THE INTERNATIONAL JOURNAL OF AVIATION PSYCHOLOGY, *14*(1), 53–77

The Efficacy of Head-Down and Head-Up Synthetic Vision Display Concepts for Retro- and Forward-Fit of Commercial Aircraft

Lawrence J. Prinzel III, J. Raymond Comstock Jr.,
Louis J. Glaab, Lynda J. Kramer, and Jarvis J. Arthur
NASA Langley Research Center
Hampton, VA

John S. Barry
Lockheed-Martin
Hampton, VA

The retrofit question concerns whether useful and effective synthetic vision displays are usable in aircraft that have limited-size display spaces. Two experiments were conducted to examine the efficacy of these displays and develop field-of-view and terrain texture recommendations for design. The first experiment examined issues of field of view and display size using an Asheville, North Carolina, synthetic vision database and fixed-based simulator. The second experiment was conducted on the NASA B-757 aircraft at Dallas/Fort Worth International Airport and investigated the efficacy of both head-down and head-up displays and generic and photorealistic terrain texture. Both experiments confirmed the retrofit capability and that all sizes and texturing methods were found to be viable candidates for synthetic vision displays. These results, future directions, and implications for meeting national aeronautic safety and capacity goals are discussed.

The Synthetic Vision Systems (SVS) element of the National Aeronautics and Space Administration's (NASA) Aviation Safety Program is striving to eliminate poor visibility as a causal factor in aircraft accidents and to increase operational ca-

Requests for reprints should be sent to Lawrence J. Prinzel, Mail Stop 152, NASA Langley Research Center, Hampton, VA 23681. E-mail: l.j.prinzel@larc.nasa.gov

pabilities of general aviation, business, and commercial aircraft. To accomplish these safety and situation awareness (SA) enhancements, the SVS concept will provide a clear view of the world ahead through the display of computer-generated imagery derived from an onboard terrain, obstacle, and airport database and enhanced vision sensor technologies.

The ability of a pilot to ascertain critical information through visual perception of the outside environment can be limited by various weather phenomena, such as rain, fog, and snow. Since the beginning of flight, the aviation industry has developed various devices to overcome these low-visibility limitations. These include attitude indicators, navigation aids, instrument landing systems (ILS), moving map displays, and terrain awareness warning systems (TAWS). All of the aircraft information display concepts developed to date, however, still require pilots to continuously perform information acquisition and decoding to update and maintain their mental model to "stay ahead" of the aircraft when outside visibility is reduced. The NASA SVS project is based on the premise that better pilot SA during low-visibility conditions can be achieved by reducing the steps required to build a mental model from disparate pieces of data through the presentation of how the outside world would look to pilots if their visibility was not restricted.

HUMAN-CENTERED SVS DISPLAYS

Although avionics have advanced significantly since Jimmy Doolittle flew the first "blind" flight in 1929, Theunissen (1993, 1997) noted that significant increases in aviation safety are unlikely to come by extrapolating from current display concepts. He further stated that, "new functionality and new technology cannot simply be layered onto previous design concepts, because the current system complexities are already too high. Better human-machine interfaces require a fundamentally new approach" (Theunissen, 1997, p. 7). Bennet and Flach (1994) argued that such an approach should not focus on development of "idiot-proof" systems because of the infinite potential problem space, but rather should provide to the pilot information that would enable successful solution sets to be generated. These displays should present continuous information about spatial constraints rather than command changes to reduce error states, and should show error margins that depict the bounds in which the pilot may safety operate in contrast to the compensatory control strategy required by current cockpit instruments. They further concluded that dynamic, graphical representations hold the greatest promise to achieve such human-centered design because they allow human flexibility to best be exploited through the presentation of natural versus coded information to the pilot.

Natural information implies the method of information acquisition by the pilot similar to that experienced in visual meteorological conditions (VMC) by looking out the window. Visual altitude judgment is an example of natural information.

Coded information implies some type of information presentation to the pilot that requires interpretation to comprehend the actual value. An example of coded information is altimeter reading. Helmetag, Kaufhold, Lenhart, and Purpus (1997) argued that it is very important to give the pilot information required to maintain SA in low-visibility conditions and that natural information presentation is intuitive and able to be perceived in a much more rapid manner than coded information. SVS displays provide exactly such a natural presentation of the outside world with proximity compatible, integrated information (Wickens & Andre, 1990) that is both intuitive and easy to process.

SAFETY BENEFITS OF SVS

Synthetic vision technology may allow the issues associated with limited visibility to be solved with a visibility-based solution, making every flight the equivalent of a clear daylight operation, which will help improve SA and support proper development of the pilots' mental model. Therefore, SVS can have a significant impact on improving aviation safety because limited visibility represents the single greatest contributing factor in many fatal airline accidents worldwide (Boeing, 1996).

Consider that one of the major types of commercial aviation accidents involving low-visibility issues is controlled flight into terrain (CFIT) and that CFIT is one of the greatest causes of aviation fatalities (Moroze & Snow, 1999). A CFIT accident is defined as "one in which an otherwise-serviceable aircraft, under control of the crew, is flown (unintentionally) into terrain, obstacles or water, with no prior awareness on the part of the crew of the impending collision" (Wiener, 1977, p. 171). A Flight Safety Foundation (FSF) analysis evinced that 90% of CFIT accidents occurred in instrument meteorological conditions and that 25% occurred with aircraft equipped with a ground proximity warning system. The FSF also reported that nonprecision approaches were five times more likely to result in a CFIT, and that lack of crew SA of terrain and aircraft position was the leading contributing cause (Khatwa & Roelen, 1998). Although TAWS may help to mitigate some of these factors, the use of the technology generally follows the warn–act model and, therefore, requires the flight crew to be reactive rather than proactive. Theoretically, TAWS provides a warning when the flight crew has already lost SA, and might not be optimal given the reaction time required to adequately recognize and assess the situation and initiate an escape maneuver (Moroze & Snow, 1999). Snow and Reising (1999) argued that what is needed is an intuitive system that improves pilot SA with respect to spatial orientation in terms of terrain and flight path, and does not require the pilot to divert visual attention and cognitive resources away from possible external events and primary flight reference. A system that can help prevent rather than just warn the flight crew of a potential collision with terrain is needed; such a system can be provided by synthetic vision.

OPERATIONAL BENEFITS OF SVS

The aviation safety benefits of synthetic vision alone are reason enough to pursue the technologies but, due to the costs associated with such a system, it must also present operational and economic benefits. NASA anticipates that SVS technology could serve to increase national airspace system capacity by providing the potential for increased VMC-type operations even under Category IIIb weather conditions (Williams et al., 2001). Benefits would include: (a) reduced runway occupancy time in low visibility; (b) reduced departure and arrival minimums; (c) better allowance for converging and circling approaches, especially for dual and triple runway configurations; (d) reduced interarrival separations; and (e) independent operations on closely spaced parallel runways. A cost–benefit analysis of 10 airports (DFW, ORD, LAX, ATL, DTW, MSP, EWR, SEA, LGA, and JFK) calculated the average cost savings to airlines for the years 2006 to 2015 to be $2.25 billion (Williams et al., 2001).

RESEARCH CHALLENGES OF SVS

Although the safety and economic advantages and payoff to pursuing SVS are great, there are significant research challenges to be addressed before SVS can be considered viable as a technological alternative. To provide a better definition of the concept of operations (CONOPS) of synthetic vision technology for commercial and business aircraft, a workshop resulting in a CONOPS document was held at the NASA Langley Research Center (Williams et al., 2001). The focus of this event was to obtain wide-ranging input from the aviation community on the benefits and features that synthetic vision might incorporate. The outcome of the workshop and subsequent activities has been the identification of numerous challenges and research issues that need to be explored in developing SVS display concepts. Many of these issues can be classified as human perceptual, such as display size and field-of-view (FOV) issues.

The issue of display size is driven largely by the need for displays compatible in size with current aircraft displays (the retrofit issue) and potential next-generation larger display surfaces (forward fit issue). Because current aircraft have either electromechanical instruments (e.g., 737–200) or small “glass” displays (e.g., 757–200), there are concerns about the efficacy of these cockpits to support SVS because of the physically smaller instrument spaces. One option to address the retrofit issue would be to present SVS on a head-up display (HUD), and research questions turn to how best to display synthetic terrain on a HUD that has limited graphical capabilities. Another option is to simply remove the traditional instruments and replace them with synthetic vision displays, and research issues then turn to whether the space constraints will allow SVS presentations to be usable by the flight crew. Because these displays have a small unity geometric FOV, the scale factor might need to be increased (i.e., minified) to allow more of the visual scene to be presented to make the SVS display effective (e.g., Roscoe, 1948). The wide-angle lens effect of increasing FOV, however, interacts with display size and can lead to perceptual distortions as the minification factor (MF) is increased (e.g., virtual space effect; McGreevy & Ellis, 1986).

There are other perceptual issues concerning the content and type of information in the pictorial scene that also need to be addressed. SVS display scenes can be constructed from terrain elevation data and smoothed with generic terrain algorithms, or they can be created by adding color and photo-realistic texture content information from aerial photographs. A research question that needs to be answered is which method provides the best information and SA gain to the pilot. Are the additional data cost and computing requirements for photorealistic terrain worthwhile in terms of enhancements to pilot performance and SA?

RESEARCH PURPOSE

Two experiments were performed to evaluate candidate FOV on each of the three display sizes on approach and landing tasks in a terrain-challenged (Asheville Airport; AVL) and a complex, nighttime operational environment (Dallas/Fort Worth International Airport; DFW). The DFW flight test also examined the efficacy of SVS presentation on a HUD. The objectives of the experiments were to address the following:

1. The FOV recommendations for head-down display (HDD) sizes (Experiments 1 & 2).
2. The effect of HDD size on pilot performance and SA enhancements (Experiments 1 & 2).
3. The effect of SVS HUD concepts on pilot performance and SA enhancements (Experiment 2).
4. The effect of generic and photo-realistic terrain texturing methods on both HUD and HDD SVS display concepts (Experiment 2).
5. The evaluation and demonstration of SVS display concepts during complex, nighttime approaches at a large international airport (Experiment 2).

EXPERIMENT 1

The objective of Experiment 1 was to examine candidate FOVs and display sizes while pilots made simulated approaches to AVL. The display sizes that were investigated were Size A, Size D, and Size X (discussed later). The FOVs for this study were unity or one-to-one, 30°, 60°, and pilot-selectable. The hypotheses for Experiment 1 included the following: (a) All display sizes would provide adequate information for the successful conduct of the approach to AVL, as determined by performance and subjective response data; and (b) there is an optimal or preferred FOV for each display size as reflected in pilot-selectable trials and subjective response data.

Method

Participants

Eight transport-rated airline captains served as test participants. Asheville was chosen from a list of domestic “terrain-challenged” airports and was the site of a 1998 NASA SVS flight research study. Figure 1 shows one synthetic vision display concept on approach to AVL that was used in Experiment 1.

FIGURE 1 Synthetic vision display concept on approach to AVL.

Simulation Facilities

The Visual Imaging Simulator for Transport Aircraft Systems (VISTAS-I) facility at the NASA Langley Research Center was used for evaluating synthetic vision display concepts for the AVL database (Figure 2). The VISTAS-I facility consists of a large HDD surface, which uses a rear projection system (2 JVC models DLA-S10U) to present the HDD concepts, and an Electrohome Marquee 8000 forward system projector to present the out-the-window (OTW) scene. Pilots were instructed to make the approaches with primary reference to the SVS display and simulated fog was used to restrict visibility and reduce pilot reliance on the OTW scene. All pilots commented that their focus was on the SVS display for the approaches until arriving at a decision height of 200 above ground level (AGL), at which time the trial was ended.

The simulation aircraft model used matched the performance capabilities of a Boeing 757-200. Pilots made approaches with manual throttles and were instructed to maintain an approach speed target of 138 kt. The AVL scene and displays were generated using Silicon Graphics Onyx-2 Infinite Reality computer, Intergraph ZX-1 dual-Pentium processor computer, and Wildcat model 4110 high-speed graphics cards. The operating system platform was Microsoft Windows NT.

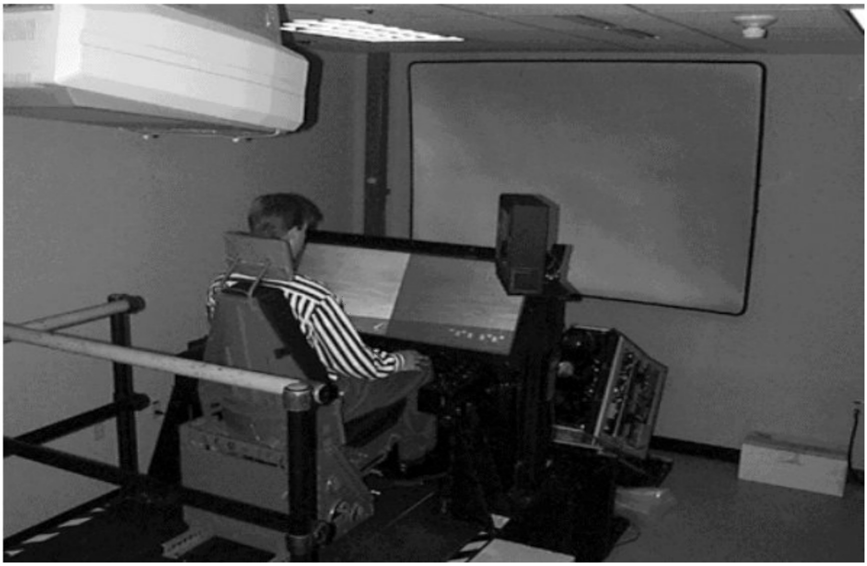

FIGURE 2 Visual imaging simulator for transport aircraft systems (VISTAS-I).

SVS Display Sizes and Format

Three display sizes were evaluated in the study and the dimensions from pilot eye reference point of these SVS display concepts are shown in Table 1. The smallest size, designated Size A, approximated the size of the electronic attitude direction indicator (EADI) in the current generation B757/B767 aircraft, and the display concept represented a retrofit concept of extracting the current EADI and replacing it with an SVS display. The A-size SVS display concept, therefore, did not incorporate airspeed, altitude, or vertical speed information, and pilots obtained the data from traditional round dials (9.5 cm in diameter) that were presented adjacent to the SVS display.

The next size represented a form factor Size D display, which approximated the size of a primary flight display (PFD) in the B777 or B747-400 aircraft. The largest of the displays tested, designated Size X, represented display sizes envisioned as a potential display size in future transport aircraft. Both the D and X displays had integrated airspeed, altitude, and vertical rate information in a moving “tape” format found in a typical PFD. Each SVS display size, including Size A, had superimposed symbology showing the horizon, body axis indicator (waterline symbol), pitch information, roll scale, horizontal and vertical path deviation scales, radar altitude (below 500 ft above ground level), and a flight path/velocity vector.

A navigation display was presented with the SVS concepts that showed moving map format waypoints (track-up) along the programmed magenta path. For Experiment 1, the SVS display showed the perspective terrain with photo-texturing of terrain features around the airport area. Photo-texturing consists of superimposing aerial photography on the terrain elevation information to re-create a realistic perspective scene. At AVL, the photo-texture covered an area 3 miles wide by 8 miles long centered about the airport. Outside the photo-textured area, generic shading of terrain features was presented.

Display FOV

Experiment 1 evaluated a subset of the possible FOVs that could be used in an SVS display. For each SVS display size, unity, 30°, and 60° FOVs were evaluated. FOV is based on horizontal FOV and vertical FOV is based on aspect ratio. For

TABLE 1
Synthetic Vision Systems Display Size Dimensions and Unity Field of View

	Size A	*Size D*	*Size X*
Width	12.9 cm	16.0 cm	25.0 cm
Height	12.6 cm	16.0 cm	20.2 cm
Horizontal	11.5°	14.2°	22.0°
Vertical	11.2°	14.2°	17.8°

75% of the experimental trials, the FOV was held fixed for each display size condition. For the remaining trials, the FOV was pilot-selectable and the pilot could change the FOV as desired at any point during the approach. Each pilot participant, therefore, was presented with each FOV option for each display size, including trials that were pilot-selectable.

Experimental Design and Procedure

A 2 (runway: 16 or 34) × 3 (display size: A, D, X) × 4 (FOV: unity, 30°, 60°, pilot-selectable) repeated experimental design was used. Display size and FOV were counterbalanced across pilot participants for a total of 12 experimental data runs. Runway was randomized for an equal number of data runs to each runway across display size and FOV. All pilots were given baseline training with a traditional EADI to familiarize them with the simulator and they participated in training runs with each display size concept before data collection began.

Six different scenarios were tested for approaches to AVL, which consisted of three starting points for the published ILS approaches to the northbound runway (RWY 34) and three starting points for the southbound runway (RWY 16). Each received an equal number of approaches to the two runways, and the six starting points were randomly presented to reduce pilot recognition and rote task completion. Each scenario began at 4,400 mean sea level (MSL) on a stabilized approach to AVL inside the initial approach fix. Data collection began 8.5 nm from runway threshold.

Using these scenarios, each test participant was presented with each factorial combination of display size and FOV option. An additional data run was performed at 90° FOV and display size was cycled to expose the pilot to the option, but no data were collected or analyzed. Performance data and subjective ratings and comments were recorded throughout the trials. After all experimental trials were completed, pilots were given a Situation Awareness Subjective Workload Dominance (SA–SWORD; Vidulich & Hughes, 1991) scale and participated in a semistructured interview and debriefing. Pilots remained seated at the simulator while completing the SA-SWORD and while participating in the semistructured interview to cycle through each display concept and FOV combination including the 90° FOV option.

Results

Pilot Performance

No significant differences were found for runway ($p > .05$) and, therefore, data were collapsed across the independent variable. For the test trials with fixed display size and FOV, there were nine combinations that could be compared at selected points on the approach. The approach segments consisted of mean path error derived over a 10,000-ft path segment. For example, the segment labeled Seg- 45 represented data obtained from –50,000 to –40,000 ft prior to runway threshold crossing. Baseline data were not collected because the objective was to evaluate the effect of display size and FOV and develop a set of recommendations. Research, however, has consistently demonstrated the advantages of SVS for pilot performance and SA. The interested reader is directed to several studies that directly compared the performance and SA benefits of SVS (Bailey, Parrish, Arthur, & Norman, 2002; Glabb & Takalu, 2002; Prinzel et al., 2002; Stark, Comstock, Prinzel, Burdette, & Scerbo, 2001; Uenking & Hughes, 2002).

Lateral path performance. A repeated measures analysis of variance (ANOVA) showed a significant effect for flight segment (Figure 3), $F(4, 28) = 9.15$, $p < .01$, but no significant main effects or interactions were reported for display size or FOV levels ($p > .05$). The significant difference for flight segment shows the effect of being more accurately on the horizontal path in proximity to the runway threshold because of the increased ILS guidance precision.

Vertical path performance. The same method for defining flight segments was used for analysis of vertical path performance. A repeated measures ANOVA also revealed a significant effect for flight segment (Figure 4), $F(4, 28) = 7.52$, $p < .01$, but no significant main effects or interactions ($p < .05$) for display size or FOV conditions. As for lateral path error, the significant difference by segment reflects decreased vertical error near the runway because of greater precision of ILS guidance nearer the runway threshold.

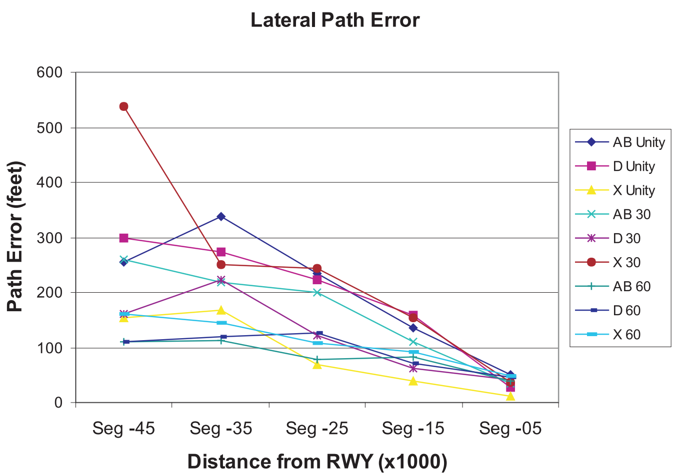

FIGURE 3 Lateral path error.

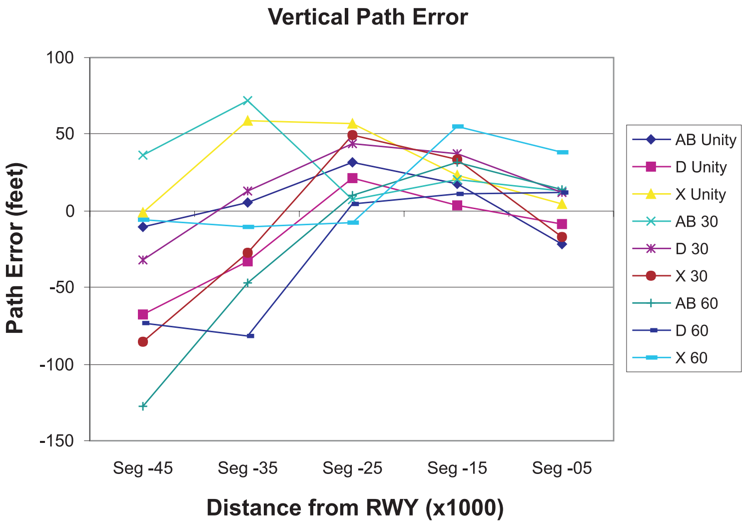

FIGURE 4 Vertical path error.

FOV Preferences

For the pilot-selectable trials, participants consistently selected a fixed FOV option approximately 4 nm to touchdown. Pilots tended to select unity (80%) and 30° FOV (15%), with only 5% of the trials being flown with a 60° FOV setting. Figure 5 shows the mean time in each FOV that mirrors these results, but also shows that the 30° FOV option was selected most often before the final approach fix, wherein pilots selected unity. Pilots rarely chose 60° or 90° FOV options except for with Size X. The distance prior to runway threshold where the last change in FOV was made was analyzed and no significant differences by display size condition were found ($p > .05$). Changes in FOV were not made near the runway and, averaging across pilot-selectable trials, the mean distance for the final FOV change was 3.7 nm prior to runway threshold crossing. It is interesting to note the decrease in larger FOV selections for the smaller display sizes, which matches subjective comments that indicated that information in the Size A display "just gets too small" with larger FOV selections.

The distribution of selected FOVs reflects pilot responses to a question asking which two FOV options they would choose if SVS was available in their cockpit. Pilots tended to choose unity (86%) and 30° (57%) FOV. Some pilots did select the 60° (28%) and 90° (28%) options because of a requirement for larger FOVs with the smaller Size A display concept (totals to 200% due to choice of two FOVs). Although 86% of pilots chose unity, it was often the second choice. Based on mean rat-

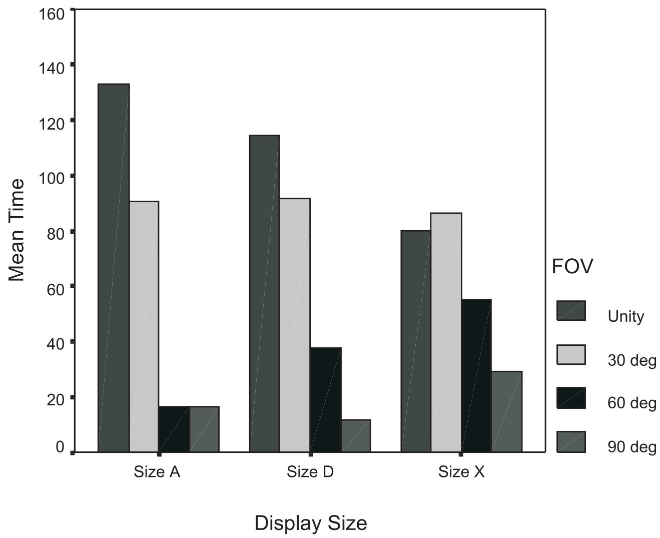

FIGURE 5 Mean seconds of FOV selection during pilot-selectable trials.

ings for order of preference, pilots preferred to have 30°, then unity, then 60°, and finally 90°. Pilots were consistent in their preferences across display sizes ($p > .05$).

SA

An SA-SWORD (Vidulich & Hughes, 1991) was administered after each block run of a display size to assess SA preferences for FOV for that size display (A, D, or X). Pilots were asked to base ratings on their operational experience and the definition of SA given as "the pilot has an integrated understanding of the factors that will contribute to the safe flying of the aircraft under normal or non-normal conditions" (Regal, Rogers, & Boucek, 1988, p. 1). Separate analyses were conducted for each display size, and ANOVAs revealed significant effects for Size A, $F(3, 18) = 131.430$, $p < .0001$; Size D, $F(3, 18) = 483.885$, $p < .0001$; and Size X, $F(3, 18) = 37.932$, $p < .001$. A Student–Newman–Keuls (SNK) post hoc analysis revealed four unique pairwise groupings: 30 > unity > 60 > 90 for Size A. For Sizes D and X, there were three unique pairwise groupings: unity > 30 > 60 = 90.

EXPERIMENT 2

The objectives of Experiment 2 were similar to those of Experiment 1 with a few exceptions. The most notable difference was that the experiment was a flight test using the NASA B-757-200 research airplane. Three different HDD configurations (Sizes A, D, and X) were evaluated during this flight test, but evaluation pilots also evaluated an SVS HUD concept. For both the HDD and HUD concepts, an evaluation of generic and photo-realistic terrain-texturing methods was also performed. The hypotheses for Experiment 2 were the following: (a) All display sizes would provide adequate information for the successful conduct of the complex, nighttime approaches to DFW; (b) there is an optimal or preferred FOV for each HDD display size; (c) the HUD would be shown to be a viable retrofit candidate; and (d) no performance differences would be found between generic and photo-realistic texture, but participants have higher preference ratings for the photo-realistic presentation.

Method

Participants

Six air-transport-rated commercial airline pilots were the participants for Experiment 2. All participants were current commercial B757 pilots who had experience with HUDs, mostly through military background. These participants were provided with familiarization training at the NASA Langley Research Center and, during the training, participated in Experiment 1.

HDD Research Display

HDD SVS concepts were presented on the SVS research display (SVS-RD), which was 14.5 in. wide by 10.9 in. tall, producing a viewing area of 158.1 square in. The display was operated in XGA mode with vertical and horizontal test resolution of 71 pixels per inch. The SVS-RD has a brightness of 900 nits and was removable in flight to address safety-of-flight concerns.

HUD Research Display

The HUD employed for Experiment 2 was a Flight Dynamics Model 2300R HGS. The FOV of the HUD was 30° horizontal by 24° vertical with a 4° look-down bias. The resulting effective FOV was 16° below and 8° above the reference waterline. Symbology and terrain information was provided to the HUD via a raster-to-stroke converter unit. Maximum brightness of the HUD image was greater than 1,000 ft.-Lamberts and brightness and contrast of both the HUD symbology and synthetic terrain was adjustable by the evaluation pilot. The evaluation pilot could view the HUD image within an eye-box approximately 5 in. wide, 2.8 in. tall, and 6 in. deep.

Display Symbology

Common symbology included a 5° increment pitch scale with reference waterline, roll scale with small tick marks every 5° and large tickmarks every 10°, bank indicator with sideslip wedge and digital magnetic heading, wind speed and relative direction, heading scale with labels every 10° and tick marks every 5°, flight path marker with acceleration along the flight path indicator, reference airspeed error, and sideslip flag. Localizer and glideslope course deviation indicators were also included. In addition, a magenta runway outline box and extended runway centerline were included for the initial runway. The navigation display (ND) included the defined path and provided primary lateral navigation guidance, prior to final approach. For the Size D and Size X SVS PFDs, airspeed, altitude, and vertical speed were presented in a nominal tape format with airspeed bugs and limit speeds present. Traditional round dials were employed for airspeed, altitude, and vertical speed for the Size A display. Airspeed and altitude were displayed digitally for the SVS-HUD concepts. Airspeed, altitude, and vertical speed were colored white on the HDDs and airspeed limits were shown in standard red and white "barber pole" format.

A minimal tunnel in the sky was incorporated into the symbology set for evaluation purposes. Intended to provide a three-dimensional representation of the intended flight path, the tunnel in the sky was presented to the evaluation pilots by magenta "crow's feet" triads located at all four corners of the defined path. The dimensions of the minimal tunnel in the sky were based on the navigation performance of standard ILS and were 1 dot wide, limited to a maximum width of 600 ft, and 2 dots high, limited to a maximum height of 350 feet and a minimum height of 50 ft. Pilots were instructed to observe the tunnel in the sky but to not use it as a guidance system nor perform closed-loop high-gain maneuvering with respect to it. The primary purpose of the tunnel in the sky was to define where the three-dimensional path was. Research has demonstrated the advantages of tunnel displays for maintaining lateral and vertical path awareness (e.g., Haskell & Wickens, 1993; Snow, Reising, Liggett, & Barry, 1999; Williams et al., 2001) and that the inclusion of synthetic terrain may significantly improve SA potential of tunnel displays (Snow & French, 2001; Snow et al., 1999; Williams et al., 2001).

Terrain Database

The DFW terrain database was generated using 1-arcsec (98 ft) postspacing digital elevation model data and covered an area of approximately 100 nm by 100 nm centered about DFW airport with an elevation accuracy of approximately 3.2 ft. One terrain-texturing option, generically textured terrain, used different color shades to represent terrain on the HDD. The HUD concept used the green RGB channel and varied color shades, rather than different colors, to reflect changes in elevation. The second terrain-texturing option was photo-realistic, and used orthorectified aerial photographs to texture the terrain to generate a highly realistic looking presentation (hence, photo-realistic) with 3 m per pixel resolution. High-resolution photo texturing was applied to an area 6 nm by 15 nm center aligned with Runways 17C and 35C. The photo-realistic HUD concept used the RGB file format and masked out the red and blue channels coming from the photo-realistic database and converted the image back to enhanced compressed wavelet (ECW) format.

Flight Task Scenarios

Four pilot tasks were employed for the DFW flight test. Two of the tasks, referred to as the straight-in approaches, required the pilot to perform a nominal downwind, baseleg, and straight-in final approach to Runways 17C and 35C. The other two tasks, referred to as the runway change or side-step maneuver, required the pilot to fly the same downwind path and initial baseleg as for the straight-in maneuvers. However, the baseleg was shortened to establish an initial final approach to either Runway 17L or 35R, depending on prevailing traffic flow at DFW. Once the aircraft was 5 nm from the initial runway, the pilots were instructed by the DFW tower to execute the side-step maneuver to runway 17C or 35C.

Figure 6 depicts the south-flow straight-in and side-step maneuver tasks for approaches to Runway 17L or 17C. All four tasks required the evaluation pilot to assume control of the aircraft abeam the midfield position of Runway 17C or 35C at 5,000 ft on downwind leg and maintain nominal approach airspeed. Just downwind of the midfield position, the pilot executed a descent to an altitude of 3,500 ft following tunnel symbology. The pilot was instructed to maintain 3,500 ft on baseleg and to execute the turn to final following the path guidance from the electronic horizontal situation indicator and tunnel symbology. Flap settings were adjusted based on nominal B-757 operations. Pilots were instructed to use the autothrottles to maintain airspeed.

For the runway change tasks, the pilot was instructed to change to Runway 17C or 35C at 5 nm from the initial runway threshold. Pilots were required to maneuver the aircraft with reference to the SVS display concept being evaluated, which also captured and presented localizer and glideslope information for the target runway.

Data Collection and Recording

Qualitative measures. Qualitative pilot ratings and comments were collected during the flight and in postflight debriefings. Pilots were also encouraged to provide a running commentary during the flight and these were recorded on the audio channel for later analysis. Responses to in-flight questions were collected once control of the aircraft was handed over to the safety pilot.

Quantitative measures. The quantitative dependent variables were root mean square (RMS) values for pilot wheel, column, and rudder pedal control in-

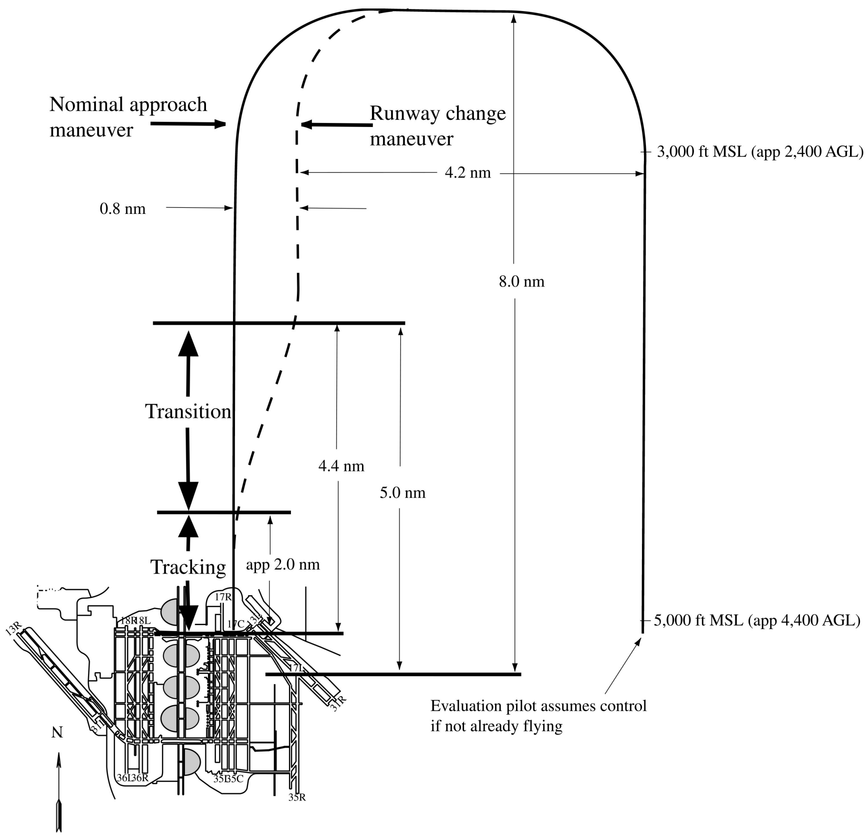

FIGURE 6 Evaluation tasks for south-flow DFW operations.

puts for workload; lateral and vertical path performance during the side-step maneuver tracking phase; maximum heading change; minimum, maximum, and mean FOV settings; and MF. MF represents the relative FOV difference between selected FOV and unity FOV for a given display size (e.g., 30° Size A = 30°/11.5° = 2.6 MF). Pilot performance data were recorded at a rate of 10 Hz and were collected on final approach 5 nm from the runway and terminated on go-around.

Data were divided into transition and tracking segments once established on final approximately 5 nm from runway threshold. The transition segment began at 5 nm from the initial runway threshold and ended when the pilot had reestablished the aircraft onto the target final approach path. RMS of bank angle, column deviation, and wheel deviation and maximum heading change were the primary dependent variables of interest for the transition phase. The tracking phase began when the pilot had reestablished the aircraft onto the target final approach path and ended at 200 ft AGL when the pilot initiated a go-around (Figure 6). Lateral and vertical path RMS error was collected and analyzed to measure pilot performance during this phase. The criteria to establish the end of the transition segment and the initiation of the tracking segment were ± 1 dot of localizer and glideslope, ± 5 degrees in track error, and ± 3° in flight path angle error.

Experimental Design and Procedure

A 4 (display type: HUD, A, X, D) × 2 (texture type: photo, generic) × 2 (runway: 17C, 35C) repeated measures, randomized experimental design was used. Runway was not randomized because runway was determined by the prevailing traffic pattern, although an equal number of runs were conducted to both ends of the runway. Experiment 2 was part of a larger flight study at DFW examining a number of aviation safety technologies that will be part of the total SVS, such as runway incursion prevention technologies. No baseline data runs were flown because of the operational, cost, and time constraints associated with combining nighttime flight tests at a busy airport and flight test objectives. The flight test took place over a 2-week period during the late evening and early morning hours when operations at DFW were fewer.

All pilots were fully briefed regarding the research objectives of the flight test, evaluation maneuvers, and data collection methods prior to each flight. Two test runs were completed with each pilot to familiarize them with the aircraft and SVS display concepts. All pilots also participated in extensive training at the NASA Langley Research Center simulation facilities.

Control was handed off to the evaluation pilot once the aircraft was established climbing in a low-workload condition. The safety pilot interacted with air traffic controllers and performed pilot-not-flying functions (e.g., air traffic control, flap settings). Once established at 5,000 ft MSL on the downwind leg, the experimental trial began and was terminated at 200 ft AGL above the runway when the evaluation pilot initiated the go-around. Once go-around checklists were completed and the aircraft was established climbing in a low-workload condition, control was transferred to the safety pilot and in-flight questionnaires were administered. After the research flight was completed, pilots participated in a semistructured interview and debriefing.

Results

Transition Phase Performance

No significant differences were found for transition or tracking phase performance for the dependent variables of RMS maximum heading error, bank angle, and column and wheel deviation ($p > .05$). An ANOVA reported comparable pilot performance for these measures regardless of display size, texture, and FOV. No differences were also found for these dependent variables between the transition phase and tracking phase (p > .05).

Tracking Phase Performance

Lateral performance. An ANOVA on the lateral tracking error during the tracking segment revealed a significant main effect for display size, $F(3, 38) = 3.10$, $p < .05$. An SNK post hoc test revealed that mean lateral path error (Figure 7) was significantly larger for Size X (112 ft) compared to HUD (49 ft), Size A (92 ft), or Size D (75 ft), which were not statistically different from each other. No significant differences were found for terrain texture, $F(1, 38) = 0.790$, p > .05, or Display × Interaction, $F(3, 38) = 1.440$, p > .05. Average RMS lateral error was 76 ft for generic and 87 ft for photo-realistic terrain texture.

Vertical performance. An ANOVA reported no significant difference for RMS vertical error across display concepts, $F(3, 38) = 0.241$, p > .05; terrain texture, $F(1, 38) = 0.378$, p > .05; or Display × Interaction, $F(3, 38) = 0.127$, p > .05. The average RMS error was 26 ft and ranged from 33 ft (Size A, photo) to 22 ft (HUD, photo). For terrain texture, average RMS vertical error was 25 ft for generic and 28 ft for photo-realistic terrain texture. Mean vertical error across display concepts is shown in Figure 7.

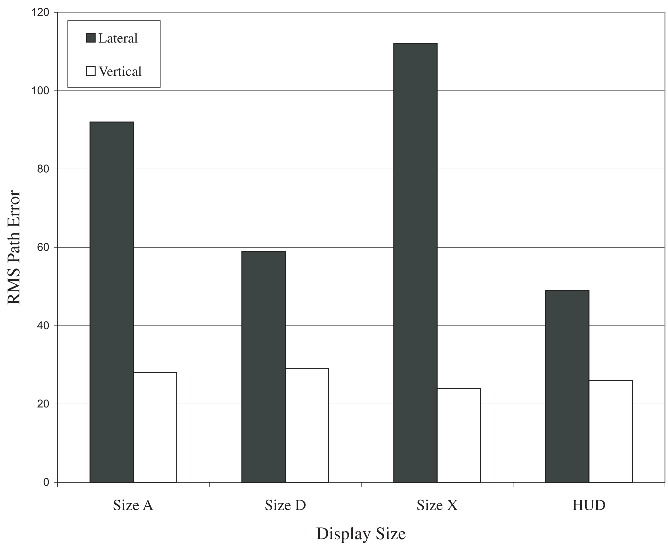

FIGURE 7 Lateral and vertical path error for tracking phase.

FOV Preferences

An ANOVA found a significant main effect for MF as a function of display size for both the transition, $F(2, 34) = 8.614$, $p < .01$, and tracking phases, $F(2, 34) = 8.146$, $p < .01$. SNK post hoc tests showed that pilots tended to chose a higher MF for Size A (3.63) than both Size D (2.44) and Size X (1.78) displays during the transition phase. For the tracking phase, pilots chose significantly higher MF for both Size A (2.39) and Size D (1.88) than for Size X (1.26). Overall, pilots selected the same FOV independent of HDD display size or terrain texture with a larger FOV during the transition phase and a smaller FOV for the tracking phase of the maneuver. Therefore, as range to touchdown decreased, the MF for the larger display sizes moved toward unity (i.e., no minification). Pilots, however, tended to select a higher minification scale factor as display size decreased because the smaller display size (i.e., Size A) was reported to be inadequate at unity FOV.

Pilot comments regarding FOV suggest that pilot-selectable FOV would be the preferred option but that the set of FOV options should be limited to a few choices to improve the ability to move quickly between the FOV modes. All pilots commented that a single FOV would not be the best solution and would impose undue restrictions on display usage. Pilots recommended that multiple FOV options based on phase of flight should be considered and all but one recommended an exclusively manual control technique for FOV selection. That single pilot suggested instead that an automatic function be implemented that changed FOV through phase of flight with a manual override capability.

Overall, a higher FOV (i.e., 50°) was recommended during early stages of an approach and smaller FOVs (e.g., 30°) were recommended for the final approach segment because of a perceived need for a smaller MF and better view of the airport environment. Pilots were asked to select two FOVs that they would select and the preferred choice was 30° and 50° FOVs, which aligns with the results from Experiment 1 (i.e., 60°) and might reflect pilot familiarity because the typical PFD provides approximately 50° (±25°) of pitch attitude.

Workload

Despite the performance data that suggest that pilots performed comparable across the HDD display concepts, pilot ratings indicated the ease of performing the approach was significantly harder with Size A, $F(2, 15) = 9.39$, $p < .01$. The scale ranged from 1 (*very hard*) to 5 (*neutral*), to 10 (*very easy*). On average, pilots gave a neutral rating to Size A (6.0), a somewhat easy rating to Size D (7.5), and a very easy rating to Size X (9.5). No significant differences were found for workload comparisons between generic and photo-realistic texture, $F(1, 20) = 3.22$, $p > .05$, and comparisons between HDD and HUD, $F(1, 20) = 0.36$, $p > .05$).

SA

An ANOVA reported no significant differences for in-flight questions regarding SA for HDD concepts, $F(2, 15) = 1.24$, $p > .05$, and ease of predicting flight path, $F(2, 15) = 1.94$, $p > .05$. Overall, as HDD size increased, maintaining SA and predicting flight path became easier, but the result was not significant. As expected, all pilots expressed the "larger is better" preference and rated Size X somewhat easy to maintain SA and predict flight path (Figure 8). No significant differences were reported between HDD and HUD, $F(1, 20) = 2.32$, $p > .05$.

A review of pilot comments indicated that only one pilot reported that the Size A display concept could not achieve an effective presentation of the synthetic terrain to significantly enhance SA compared to an EADI. All pilots, however, noted that large MFs (e.g., Size A at 60° FOV) produced an illusion that objects in the SVS scene were much farther away and that perceived altitudes were lower than actual. Larger MFs (i.e., > 4.8) also created significant runway viewing problems because objects subtended at angles on the display smaller than in the real world.

Another question asked of pilots was their preferences for photo-realistic or generic texture (Figure 8). Despite pilot performance results that showed no differences between the two texture methods, pilot ratings indicated that it was easier to maintain SA with photo-realistic than generic texturing, although it was not found to be significant, $F(1, 20) = 2.54$, $p > .05$. These pilots were very familiar with the Dallas-Fort Worth area and noted the depiction of shopping malls, roads, and population areas was very helpful in maintaining SA. Other comments included that the photo-realistic texture helped determine rate of closure with objects on the

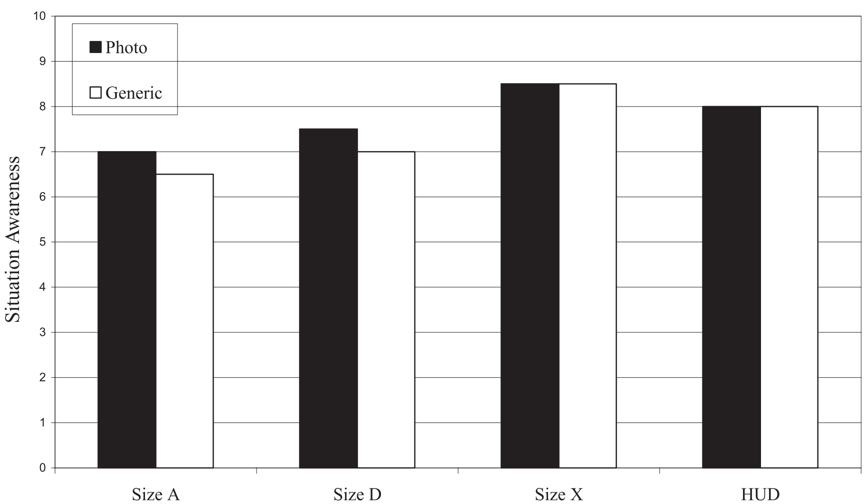

FIGURE 8 Pilot ratings of situation awareness for display size and texture.

ground and supplied cuing for runway centerline alignment. Generic texturing, however, was thought to be better for nonterminal operations because the level of detail found in photo-realism would not be necessary and cultural features stood out better against the generic terrain.

DISCUSSION

Synthetic vision has the potential to provide significant safety and economic benefits, particularly if the system is effective as both a retrofit and forward-fit solution to visibility-restricted problems. Previous research has shown the efficacy of synthetic vision on large-size displays and, therefore, synthetic vision is expected to be capable of effective presentation as glass displays become larger with each generation of aircraft. Because the majority of the current commercial aircraft fleet has electromechanical instruments or limited glass real estate, however, any significant benefits would require answering the retrofit question of whether effective presentation of synthetic vision can also be made in current aircraft cockpits.

Display Size

The retrofit question concerns our hypothesis that the HUD and smaller SVS display sizes would provide adequate information to enable the pilot to make safe and precise approaches. The results of the experiments confirmed the hypothesis and suggest that SVS is viable as a retrofit candidate. Experiment 1 showed no differences in path performance between display sizes or FOV, and Experiment 2 showed differences only for the HUD concept for lateral path performance.

One explanation for the superiority of the HUD is that the most frequently selected FOV was 30° for all the HDD concepts, which represents an MF of 2.67 for Size A, 1.82 for Size D, 1.31 for Size X, and only 1.00 for HUD (unity FOV is 30°). The pilots would use the flight path marker to center on the runway in the synthetic scene to shoot the approach. However, as the MF was increased, greater path error was required to displace the flight path marker to be noticeable to the pilot. Despite this, the difference between the worst and best lateral performance was approximately 61 ft from the runway threshold during final approach. Considering that there were also no differences found in vertical path performance, the result can be interpreted as being not practically significant.

FOV

Another hypothesis of the experiments was that there was an optimal or preferred FOV setting for synthetic vision displays, and this was confirmed by the results of the experiments. The SA-SWORD and pilot preference data from Experiment 1 showed that pilots preferred 30° and unity, and pilots in both experiments used these FOV options between 90% (Size A) and 70% (Size X) of the approach. Overall, the option of 90° was found unusable for both Size A and Size D because of the high MF (8.00 and 5.48, respectively), and was used only 12% of the time with the Size X display. Pilot comments noted that 90° was difficult to use on approach because of the precision required, but might be optimal for the en route phase of flight in which the increased visual scene would help with SA.

Terrain Texture

The final hypothesis proposed that pilots would prefer the photo-realistic terrain presentation over the generic terrain texture, but would reveal no differences in pilot performance. The results of Experiment 2 confirm this hypothesis. No significant main effects or interactions were found for pilot performance as a function of terrain texture. The mean difference between generic and photo-realistic was 15.5 ft lateral and 2.6 ft vertical during the tracking phase. All pilots, however, commented and gave higher subjective ratings to the photo-realistic concept, and this represents a common dissociation in display evaluation, where participants prefer one concept to another but show no differences in performance.

PRACTICAL APPLICATION

Taken together, the conclusions that can be drawn from these experiments are that synthetic vision can be implemented on retrofit sizes and, therefore, can successfully be introduced into the current aircraft fleet. To be effective, synthetic vision presented on small display sizes would have to be minified and our results indicate that the MF should not exceed 4.5 for optimal performance, although more research is needed to confirm such a conclusion.

Because no performance differences were found between photo-realistic and generic terrain texture methods, the generic terrain presentation might represent an effective and lower cost option for synthetic vision displays, although photo-realistic terrain does have properties that can increase the margin for safety and operations. Several pilots did comment that photo-realistic texture would be helpful for SA during climb, en route, and descent phases of flight. A recent flight test in the terrain-challenged area of Eagle-Vail, CO, however, found no performance or SA penalties for the generic texture concept, although pilots reported an overall preference for the photo-realistic presentation (Bailey et al., 2002; Prinzel et al., 2002). The NASA Aviation Safety Program is currently evaluating a synthetic vision concept that combines generic and photo-realistic terrain texture to take advantage of the benefits both methods offer for SA.

FUTURE DIRECTIONS

The problem of reduced visibility challenges aviation goals to reduce the accident rate and improve operational capacity (Federal Aviation Administration, 2001; NASA, 2001). The approach of synthetic vision is to solve the problem through the presentation of how the outside world would look to the pilot if vision were not restricted. TAWS are steps in the right direction and they have significantly improved safety, but the solution treats the symptoms and not the cause (Moroze & Snow, 1999). Synthetic vision instead provides for proactive prevention of visibility-induced accidents while also increasing the capability to make approaches in weather conditions and airports not currently legal for low-visibility operations. Although our research did not specifically address these aviation safety and operational benefits, subsequent studies (e.g., Prinzel et al., 2002) have substantiated the performance and SA enhancements of synthetic vision even while making complex, circling approaches under conditions that are beyond current cockpit technology capabilities. Furthermore, the concept described here represents only the database and display concepts and not the total SVS, which will include synthetic vision navigation displays; runway incursion prevention technology; database integrity monitoring equipment; enhanced vision sensors; taxi navigation displays; and advanced communication, navigation, and surveillance technologies (McCann et al., 1998; Timmerman, 2001; Uijt de Haag, Young, Sayre, Campbell, &Vadlamani, 2002; Williams et al., 2001; Young & Jones, 2001). These technologies represent a comprehensive solution that will be evaluated in near-term NASA simulation and flight research. Together, synthetic vision may considerably help meet national aeronautic goals to "reduce the fatal accident rate by a factor of 5" and to "double the capacity of the aviation system," both with 10 years (NASA, 2001, p. 2).

REFERENCES

Bailey, R. E., Parrish, R. V, Arthur, J. J., & Norman, R. M. (2002). Flight test evaluation of tactical synthetic vision display concepts in a terrain-challenged operating environment. In J. G. Verly (Ed.), *Enhanced and synthetic vision 2002* (pp. 178–189). Bellingham, WA: International Society for Optical Engineering.

Bennet, K. B., & Flach, J. M. (1994). When automation fails …. In M. Mouloua & R. Parasuraman (Eds.), *Human performance in automated systems: Current research and trends* (pp. 229–234). Mahwah, NJ: Lawrence Erlbaum Associates, Inc.

Boeing. (1996). *Statistical summary of commercial jet aircraft accidents, worldwide operations, 1959–1995.* Seattle, WA: Airplane Safety Engineering, Boeing Commercial Airplane Group.

Federal Aviation Administration. (2001). *FAA strategic plan.* Washington, DC: Author.

Glabb, L. J., & Takalu, M. A. (2002). *Preliminary effect of synthetic vision systems displays to reduce low-visibility loss of control and controlled flight into terrain accidents* (SAE Tech. Paper No. 2002-01-1550). Warrendale, PA: Society of Automotive Engineers.

Haskell, I. D., & Wickens, C. D. (1993). Two- and three-dimensional displays for aviation: A theoretical and empirical comparison. *International Journal of Aviation Psychology, 3,* 87–109.

Helmetag, A., Kaufhold, R., Lenhart, P., & Purpus, M. (1997). *Analysis, design, and evaluation of a 3D flight guidance display.* Kyoto, Japan: Institute of Flight Mechanics and Control.

Khatwa, R., & Roelen, A. (1998, November). An analysis of controlled-flight-into-terrain (CFIT) accidents of commercial operators, 1988 through 1994. *Flight Safety Digest,* pp. 166–212.

McCann, R. S., Foyle, D. C., Hooey, B. L., Andre, A. D., Parke, B., & Kanki, B. (1998). *An evaluation of the Taxiway Navigation and Situation Awareness (T-NASA) system in high-fidelity simulation* (SAE Tech. Paper No. 985541). Warrendale, PA: Society of Automotive Engineers.

McGreevy, M. W., & Ellis, S. R. (1986). The effect of perspective geometry on judged direction in spatial information instruments. *Human Factors, 28,* 439–456.

Moroze, M. L., & Snow, M. P. (1999). Causes and remedies of controlled flight into terrain (CFIT) in military and civil aviation. In R. S. Jensen (Ed.), *Proceedings of the 10th International Symposium on Aviation Psychology.* Columbus: Ohio State University.

National Aeronautics and Space Administration. (2001). *Aerospace technology enterprise.* Washington, DC: Author.

Prinzel, L. J., Kramer, L. J., Comstock, J. R., Bailey, R. E., Hughes, M. F., & Parrish, R. V. (2002). NASA synthetic vision EGE flight test. *Proceedings of the Annual Human Factors and Ergonomics Meeting, 46,* 135–139.

Regal, D. M., Rogers, W. H., & Boucek, G. P. (1988). *Situational awareness in the commercial flight deck: Definition, measurement, and enhancement* (SAE Tech. Paper No. 881508). Warrendale, PA: Society of Automotive Engineers.

Roscoe, S. N. (1948). The effects of eliminating binocular and peripheral monocular visual cues upon airplane pilot performance in landing. *Journal of Applied Psychology, 32,* 649–662.

Simon, H. A. (1981). *The sciences of the artificial.* Cambridge, MA: MIT Press.

Snow, M. P., & French, G. A. (2001). *Human factors in head-up synthetic vision display* (SAE Tech. Paper No. 2001-01-2652). Warrendale, PA: Society of Automotive Engineers.

Snow, M. P., & Reising, J. M. (1999). Effect of pathway-in-the-sky and synthetic terrain imagery on situation awareness in a simulated low-level ingress scenario. In *Proceedings of the 4th Annual Symposium on Situation Awareness in the Tactical Air Environment* (pp. 198–207). Patuxent River, MD: Naval Air Warfare Center Aircraft Division.

Snow, M. P., Reising, J. M., Liggett, K. K., & Barry, T. P. (1999). Flying complex approaches using a head-up display: Effects of visibility and display type. In R. S. Jensen (Ed.), *Proceedings of the 10th International Symposium on Aviation Psychology.* Columbus: Ohio State University.

Stark, J., Comstock, J. R., Prinzel, L. J., Burdette, D., & Scerbo, M. W. (2001). A preliminary examination of situation awareness and pilot performance in a synthetic vision environment. *Proceedings of the Human Factors & Ergonomics Society, 45,* 40–43.

Theunissen, E. (1993). A primary flight display for four-dimensional guidance and navigation—Influence of tunnel size and level of additional information on pilot performance and control behavior. In *Proceedings of the AIAA Flight Simulation Technologies Conference* (pp. 140–146). Scottsdale, AZ: AIAA.

Theunissen, E. (1997). *Integrated design of a man–machine interface for 4-D navigation.* Delft, Netherlands: Delft University Press.

Timmerman, J. (2001). *Runway incursion prevention system—ADS-B and DGPS data link analysis, Dallas–Ft. Worth International Airport* (Tech. Rep. No. NASA CR-2001-211242). Hampton, VA: NASA Langley Research Center.

Uenking, M. D., & Hughes, M. F. (2002). *The efficacy of using synthetic vision terrain-textured images to improve pilot situation awareness* (SAE Tech. Paper 2002-01-2970). Warrendale, PA: Society of Automotive Engineers.

Uijt de Haag, M., Young, S., Sayre, J., Campbell, J., & Vadlamani, A. (2002). DEM integrity monitor equipment (DIME) flight test results. In J. G. Verly (Ed.), *Enhanced and synthetic vision 2002* (pp. 72–83). Bellingham, WA: International Society for Optical Engineering.

Vidulich, M. A., & Hughes, E. R. (1991). Testing a subjective metric of situation awareness. *Proceedings of the Human Factors & Ergonomics Society, 35,* 1307–1311.

Wickens, C. D., & Andre, A. D. (1990). Proximity compatibility principle and information display: Effects of color, space, and objectness on information integration. *Human Factors, 32,* 61–77.

Wiener, E. L. (1977). Controlled flight into terrain accidents: System-induced errors. *Human Factors, 19,* 171–181.

Williams, D., Waller, M., Koelling, J., Burdette, D., Doyle, T., Capron, W., et al. (2001). *Concept of operations for commercial and business aircraft synthetic vision systems* (NASA Tech. Memo. No. TM-2001-211058). Langley, VA: NASA Langley Research Center.

Young, S., & Jones, D. R. (2001). Runway incursion prevention: A technology solution. *Proceedings of the Flight Safety Foundation Air Safety Seminar, 54,* 1–18.

Manuscript first received March 2002